

# 経営哲学に関する講演録を用いた質問応答システム

三品 博崇<sup>†</sup> 青山 敦<sup>‡</sup> 前田 亮<sup>†</sup>

<sup>†</sup>立命館大学情報理工学部 〒525-0058 滋賀県草津市野路東1丁目1-1

<sup>‡</sup>立命館大学テクノロジー・マネジメント研究科 〒567-8570 大阪府茨木市岩倉町2-150

E-mail: <sup>†</sup>is0281vs@ed.ritsumeai.ac.jp, <sup>‡</sup>aoyama@mot.ritsumeai.ac.jp <sup>†</sup>amaeda@is.ritsumeai.ac.jp

**あらまし** 近年、大企業を含む様々な企業で不正などの発覚により大きな損害を被ったり信頼をなくす企業が増えてきている。これらのことが起こるのは企業のルールや経営に対する理念や哲学が欠落していることから起こることだと考えられる。こういった現状を踏まえ、どのように経営をしていくべきかを解決するために、本論文では、稲盛和夫氏の盛和塾での講演録をテキスト化したものを用いて、質問者の質問を解析して、質問に対して稲盛和夫氏ならばこう答えるであろう回答を、講演録テキストの中から抽出し出力する質問応答システムを提案する。

**キーワード** 稲盛和夫, word2vec, non-factoid 型, 経営哲学

## 1. はじめに

近年、大企業を含む様々な企業で不正などの発覚により大きな損害を被ったり信頼をなくす企業が増えてきている。これらのことが起こるのは企業のきちんとしたルールや経営に対する理念や哲学が欠落していることから起こることだと考えられる。また、経営者としてまだ未熟な人には、どのように経営をして行かなければならないのかを考えることは難しいと思われる。こういったことに関する方法や疑問などを質問応答形式にすることによりユーザの必要とする情報を入手できると考えられる。

質問応答では、factoid 型と non-factoid 型の二つがあり、factoid 型は名称や日付・数値など事実に基づく正答を求める質問である。non-factoid 型は理由や事象の説明に基づく正答を求める質問である。non-factoid 型質問応答でよく扱われる質問は「～はどんな人ですか?」「～とはなんですか」という質問に答える定義型質問応答、「なぜ～ですか?」「～の原因はなんですか」という質問の Why 型質問応答、「～はどうですか」「どうすれば～できますか?」という How 型質問応答の3種類である。factoid 型、non-factoid 型それぞれにおいて研究が盛んであり、多くのシステムも開発されている。しかし、一般的な質問応答システムは事実による回答を抽出するものがほとんどであり、特定の人物の考えをベースとした質問応答システムはあまり研究がされていない。これは、特定の人の考えなどを考慮するには、その人物の喋り方やよく使用する単語などを分析する必要があり、またその人物の会話データなどを多く収集する必要があるため実現するのが困難だったのではないかと考えられる。

経営に対する考え方に関して、近年では、経営哲学

というものが経営者の中で関心が出てきており、経営哲学をベースにして経営をする経営者が増えてきている。

経営哲学とは経営の手法を哲学と織り交ぜたものであり、京セラ創始者である稲盛和夫氏による経営手法である。特に経営哲学は自分のことしか考えないのではなく相手のことも考えるという「利他の心」をベースとした考え方である。現在の多くの経営者がこの経営哲学を学ぶ為に、稲盛和夫氏本人が経営者の質問に答える盛和塾に参加している。この経営哲学によって稲盛和夫氏は会社を大きくすることに成功しており、また企業の再建にも貢献をしている。経営哲学に関する質問応答を実装することによって、その都度の状況に応じて質問をすることが可能になり、その回答を参考にすることによって不正や損害などを防ぐことができると思われる。

以上から稲盛和夫氏の経営哲学に関する質問応答システムを構築する。

## 2. 関連研究

特定の人物の考えを回答とする質問応答研究は著者の知る限り存在しないと考えられる。ここでは non-factoid 型及び Why 型質問応答についての研究を述べる。

渋沢ら[1]は Why 型質問を表す文と、その回答を表す文の、文章内における位置関係に着目した Why 型質問応答システム RE:Why を実装した。この研究では質問者は検索語を入力し、検索語のみ、検索語+「なぜ」、検索語+「どうして」の3種類のクエリから Google 検索により文書を取得する。取得した文書に対して渋沢らが作成した理由語、前方指示語、後方指示語、疑問語を質問に対する回答として抽出する特徴語とし、

これらの特徴語を用いて、文書中から質問に対する回答となりうる範囲を特定する。本手法の精度は Google の最大精度が 41.5% に対して 60.7% と精度が 19.2% 向上している。本研究では回答候補抽出でこの手法を使用する。

車ら[2] は文章構造を用いた Why 型質問応答システムを提案している。車らは回答候補検索で得られた回答候補に対して Personalized PageRank 手法によって回答候補のランキングをしている。Personalized PageRank [3] とは文書の重要文抽出でよく用いられる手法であり、PageRank の Random Surfer Model に優先すべきノードにジャンプするテレポーション確率を導入したものであり、既存手法に比べて Personalized PageRank 手法の方が性能が向上した。

### 3. 提案手法

本節では、提案手法について述べる。本研究では、Why 型、How 型、What 型の 3 種類の質問に対応する質問応答システムを目指す。本論文では Why 型の質問応答システムの提案手法を述べる。Why 型質問応答システムの概要を図 1 に示す。質問文は質問者が入力した文であり、通常の質問応答システムと同じように質問文解析、文書検索、回答候補抽出、回答選択の流れで処理をする。

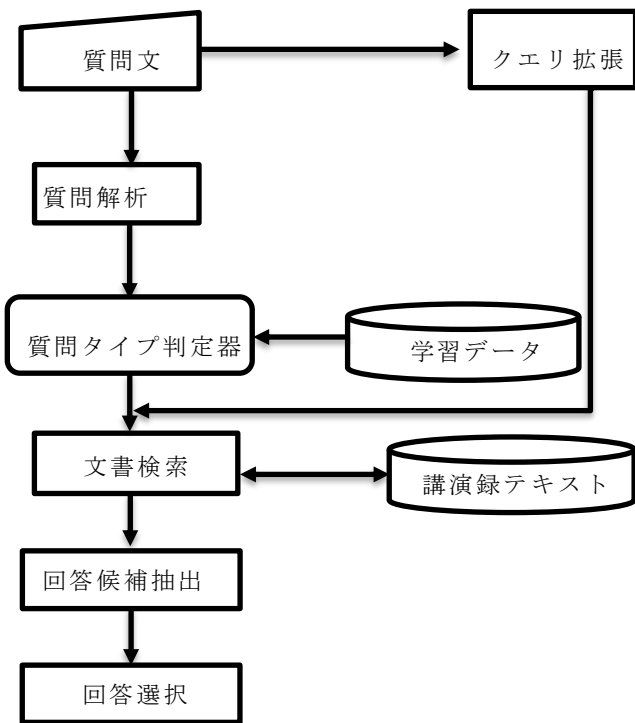


図 1 提案手法概要

### 3.1 質問解析

質問解析では質問タイプとして「Why 型」、「How 型」、「What 型」の 3 種類に設定する。また質問タイプの判定手法として機械学習アルゴリズム Support Vector Machine (SVM) を使用する。

#### 3.1.1 学習データ

本研究では、質問解析に使用する学習データとして Yahoo 知恵袋 68 件を手で取得し、データとして使用した。Yahoo 知恵袋 68 件の内訳とラベルを表 1 に示す。

表 1 ラベルと件数

質問タイプ	ラベル	件数
Why 型	1	19
How 型	2	26
What 型	3	23

#### 3.1.2 SVM (Support Vector Machine)

本研究では、機械学習ライブラリとして Scikit-learn を用いる。Support Vector Machine (SVM) は Vapnik らによって提唱された機械学習手法であり、正例と負例の 2 クラス間の距離を最大にするという基準で分離平面を決定する。本研究では、形態素解析器 MeCab によって文書から品詞が副詞の単語を抽出し、それを学習データの素性とする。これは、Why 型や How 型を判定する上で重要な単語は「なぜ」や「どうして」という単語であり、名詞や動詞などは質問タイプ判定には重要ではないと考えたからである。また、SVM に使用する特徴ベクトルとして Bag of Words (BoW) を使用する。BoW とは各文章に特徴語が何個あるかカウントして特徴ベクトルを作る手法である。

### 3.2 文書検索

本研究では、文書検索として TF-IDF 法を用いる。本研究では質問文から抽出した名詞に対して word2vec によるクエリ拡張を行う。word2vec [4][5] はニューラルネットワークを用いた単語の分散表現であり、単語に対する意味的な計算を可能とする。word2vec によって追加したクエリを含む各クエリの文書に対する TF-IDF 値をそれぞれ計算し、ある文書に対する各クエリの TF-IDF 値の和をとり、値が高い上位 20 件の文書を回答候補文書とする。また、質問文から抽出したクエリにおいて一つでも TF-IDF 値が 0 のクエリがあった場合、その文書は回答候補文書から除外する。これは質問文から抽出したクエリがその文書に存在しないので、質問に対する回答をその文書が含んでいる可能性が低いと考えたからである。

#### 3.2.1 講演録テキスト

ここでは、講演録テキストについて述べる。講演録

テキストとは稲盛和夫氏が盛和塾などで講演した内容をテキスト化したものである。経営哲学において重要となる「利他」関連の内容を表す文に対して人手で<rita>タグが付与されている。本研究ではこの<rita>タグが付与された文だけを集めた講演録テキストを文書検索で使用する。

### 3.3 回答候補抽出

3.2節で取得した上位20件の回答候補文書に対して、各文書の文中から質問に対する適切な回答を抽出する。文中から回答になりうる文を抽出するために疑問語や前方指示語を含む文を特定し、特徴語に応じて前後の文とともに抽出する。

#### 3.3.1 特徴語

本実験では Why 型の質問応答において疑問語などの特徴語を活用する。特徴語に関しては渋谷ら[1]の RE:Why システムに使われている特徴語を参考にした。特徴語の種類と特徴語の一覧の一部を表2に示す。前方指示語かつ理由語が含まれる文があった場合、理由や原因の文がその文よりも上にあると仮定する。例を図2に示す。

表2 特徴語の種類と一覧

特徴語の種類	特徴語の一覧
疑問語	なぜ、何故、どうして、？
前方指示語 かつ理由語	だから、ですから、したがって、ゆえに、すると
理由語	理由、原因、ため、訳
前方指示語	それ、これ、そういった、以上
後方指示語	以下、後述、示す、述べる

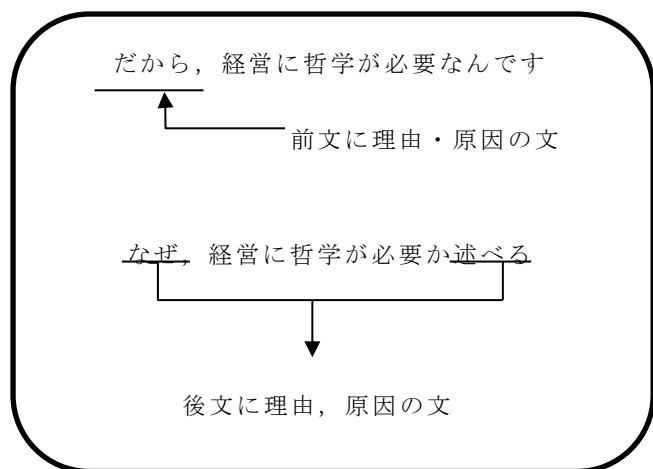


図2 原因・理由文特定の例

### 3.4 回答選択

この節では、3.3節で抽出した回答候補と質問文との類似度を計算してスコアが高い順に並べ替え、回答を出力する手法について述べる。本研究では、類似度計算として、質問文の名詞のベクトルと回答候補の名詞のベクトル同士のコサイン類似度とする。回答出力に関して、本研究では稲盛和夫氏の考えを模倣した質問応答システムであり、必ずしも正解が一つであるとは限らず、また同じ質問でも質問者の現在の状況や考えによっては質問者に最も有益な回答も変わると考えられるので、本研究の回答出力は最もスコアが高かった回答だけを出力するのではなく、ランキング形式として回答を出力する手法にする。

## 4. 評価・実験

3節で説明した提案手法を実装し、本研究の評価実験を行なった。word2vec で使用する学習データと文書検索の検索対象は講演録テキストとした。以下に実験結果と内容について述べる。

### 4.1 実験結果

本研究では質問文入力から出力する回答を20件とする。回答が質問に対して適した回答かは第一著者の判断とした。20件の回答のうち、質問に対して適した回答であると思われる回答を表3に示す。また回答自体が長い場合、回答記載は一部省略して記載している。

## 5 考察

本論文では、Why 型を対象として実験を行った。3件の質問と回答をみると、どの質問にも4位の回答が質問に適した回答となっている。また10位以降の回答の方が上位の回答よりも質問に対する回答としてふさわしいものがいくつか見受けられた。これは今回の実験では「利他の心」などの稲盛和夫氏に特有の言葉に対して特に特別な重み付けや処理を行っておらず単純にコサイン類似度計算をしているために上位に適した回答がこないと考えられる。さらに今回使用した講演録テキストに関して、文書によって文の長さが10文程度のものであれば、100文程度の長さなど様々であり、文が多い文書ほど、ある単語の出現回数が高くなり、結果的に文が多い文書ばかりが回答として出力されてしまい、文が短かくても回答に適している文書を文書検索の時点で取得できなかった可能性がある。よって短い文書に対しても取得ができるようにすればより性能が向上すると思われる。稲盛和夫氏の特有の単語を講演録テキストから分析をし、利他関連の単語及び利他関連の文でよく使われている単語に対して重みを他の単語よりも大きくすることにより精度が上がると思われる。

表3 評価実験における質問と回答の一部

質問	なぜ経営に利他が必要なのですか
1位	良い結果を期待するのなら、相手に良かれしと願い....つまり利他の心、思いやりの心を持って事業を行うことが必要不可欠なのです
4位	私心に塗れた考えで物事を決めていったのでは、社会に対しても害毒を流しますし、働く従業員たちにも大変な迷惑をかけてしまうことになります
11位	強い利己は事業をする上で不可欠なものだからです。...利他を出してくるようにしなければなりません

質問	なぜ社員がうまくまとまらないのでしょうか
4位	経営でも、美しい心でみんなのために一生懸命頑張るのであれば....喜んで働けるような気持ちにならなければ、会社というものはうまくいきません
10位	「より良い仕事をする」という節の最初に出てきますのは....みんなのために努力を惜しまなかったからこそ素晴らしい集団を築くことができたのです

質問	なんで利己で経営をしてはいけないのか？
4位	我々の判断いかんによって、企業の盛衰は...その判断が利己だけであれば、それは差はつかないでしょう
10位	私が皆さんに利他の心の大切さを説いているのは...そしてやがて社会の支持を得られなくなり、没落を遂げていく。そういうケースが多いのです

## 6 まとめ

本研究では、経営哲学に関する質問応答システムとして質問解析、文書検索、回答候補抽出、回答選択について手法を提案した。今回の実験では質問文と回答候補との類似度計算をして回答を出力していたが、講演録テキストの内容には、「なぜ経営に哲学が必要か」という質問に直接直結するタイトルの講演録文書がいくつか見受けられた。タイトルはその文書の内容をそのまま表した文なので、こういった場合には類似度計算を行わず、そのまま取得すればさらに性能が向上すると考えられる。また、本研究の最終目標は Why 型、How 型、What 型の 3 種類に対応した質問応答システムであるが、今回は Why 型の手法のみを提案した。今後 How 型、What 型を研究し、対応できる質問を広げることでより性能の良い質問応答システムができると考えられる。

## 謝 辞

本研究にあたっては、京セラ株式会社のフィロソフィ教育推進部経営研究部責任者である木谷重幸氏から提供していただいた稲盛和夫氏の講演録テキストを利用させていただき、また情報の提供や助言を賜りました。ここに深甚なる謝意を表します。

## 参 考 文 献

- [1] 洪沢潮, 林貴宏, 尾内理紀夫“Why 型質問の回答文を Web ページから抽出するシステム RE:Why の

試作”, コンピュータソフトウェア, 48(3), pp. 1512-1523 (2007)

- [2] 車智修, 鍋島啓太, 水野淳太, 岡崎直観, 乾健太郎”文章構造を用いた Why 型質問応答システム”, JSAI2013, 3B1-1(2013)
- [3] Glen Jeh and Jennifer Widom, Scaling personalized web search, in Proceedings of the 12<sup>th</sup> international conference on World Wide Web, pp. 271-279 (2003)
- [4] word2vec, <https://code.google.com/p/word2vec/>
- [5] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg Corrado, Jeff Dean, Distributed representations of words and phrases and their compositionality, (2013)